Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations

Okada, Yukinori ; Sim, Xueling ; Go, Min Jin ; Wu, Jer-Yuarn ; Gu, Dongfeng ; Takeuchi, Fumihiko ; Takahashi, Atsushi ; Maeda, Shiro ; Tsunoda, Tatsuhiko ; Chen, Peng ; Lim, Su-Chi ; Wong, Tien-Yin ; Liu, Jianjun ; Young, Terri L ; Aung, Tin ; Seielstad, Mark ; Teo, Yik-Ying ; Kim, Young Jin ; Lee, Jong-Young ; Han, Bok-Ghee ; Kang, Daehee ; Chen, Chien-Hsiun ; Tsai, Fuu-Jen ; Chang, Li-Ching ; Fann, S-J Cathy ; Mei, Hao ; Rao, Dabeeru C ; Hixson, James E ; Chen, Shufeng ; Katsuya, Tomohiro ; Isono, Masato ; Ogihara, Toshio ; Chambers, John C ; Zhang, Weihua ; Kooner, Jaspal S ; Albrecht, Eva ; Yamamoto, Kazuhiko ; Kubo, Michiaki ; Nakamura, Yusuke ; Kamatani, Naoyuki ; Kato, Norihiro ; He, Jiang ; Chen, Yuan-Tsong ; Cho, Yoon Shin ; Tai, E-Shyong ; Tanaka, Toshihiro ; Devuyst, Olivier

Abstract: Chronic kidney disease (CKD), impairment of kidney function, is a serious public health problem, and the assessment of genetic factors influencing kidney function has substantial clinical relevance. Here, we report a meta-analysis of genome-wide association studies for kidney function-related traits, including 71,149 east Asian individuals from 18 studies in 11 population-, hospital- or family-based cohorts, conducted as part of the Asian Genetic Epidemiology Network (AGEN). Our meta-analysis identified 17 loci newly associated with kidney function-related traits, including the concentrations of blood urea nitrogen, uric acid and serum creatinine and estimated glomerular filtration rate based on serum creatinine levels (eGFRcrea) ($P < 5.0 \times 10(-8)$). We further examined these loci with in silico replication in individuals of European ancestry from the KidneyGen, CKDGen and GUGC consortia, including a combined total of 110,347 individuals. We identify pleiotropic associations among these loci with kidney function-related traits and risk of CKD. These findings provide new insights into the genetics of kidney function.

DOI: https://doi.org/10.1038/ng.2352

multiple loci associated with kidney function-related traits in east Asian populations. Nature Genetics, 44(8):904-909.
DOI: https://doi.org/10.1038/ng.2352

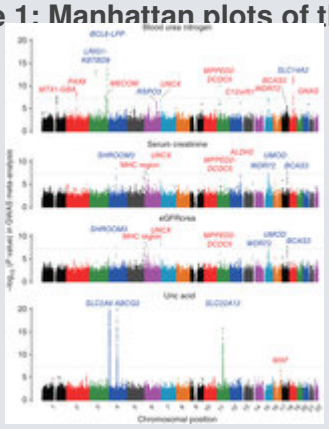
Meta-analysis identifies multiple loci associated with kidney function–related traits in east Asian populations

Yukinori Okada,[1, 2, 41]Xueling Sim,[3, 4, 41]Min Jin Go,[5, 41]Jer-Yuarn Wu,[6, 7, 41]Dongfeng Gu,[8, 41]Fumihiko Takeuchi,[9, 41]Atsushi Takahashi,[1]Shiro Maeda,[10]Tatsuhiko Tsunoda,[11]Peng Chen,[12]Su-Chi Lim,[12, 13, 14]Tien-Yin Wong,[15, 16, 17]Jianjun Liu,[18]Terri L Young,[19]Tin Aung,[15, 16]Mark Seielstad,[20]Yik-Ying Teo,[3, 12, 18, 21, 22]Young Jin Kim,[5]Jong-Young Lee,[5]Bok-Ghee Han,[5]Daehee Kang,[23]Chien-Hsiun Chen,[6, 7]Fuu-Jen Tsai,[7]Li-Ching Chang,[6]S-J Cathy Fann[6]*et al.*

Chronic kidney disease (CKD), impairment of kidney function, is a serious public health problem, and the assessment of genetic factors influencing kidney function has substantial clinical relevance. Here, we report a meta-analysis of genome-wide association studies for kidney function–related traits, including 71,149 east Asian individuals from 18 studies in 11 population-, hospital- or family-based cohorts, conducted as part of the Asian Genetic Epidemiology Network (AGEN). Our meta-analysis identified 17 loci newly associated with kidney function–related traits, including the concentrations of blood urea nitrogen, uric acid and serum creatinine and estimated glomerular filtration rate based on serum creatinine levels (eGFRcrea) ($P < 5.0 \times 10^{-8}$). We further examined these loci with *in silico* replication in individuals of European ancestry from the KidneyGen, CKDGen and GUGC consortia, including a combined total of ~110,347 individuals. We identify pleiotropic associations among these loci with kidney function–related traits and risk of CKD. These findings provide new insights into the genetics of kidney function.

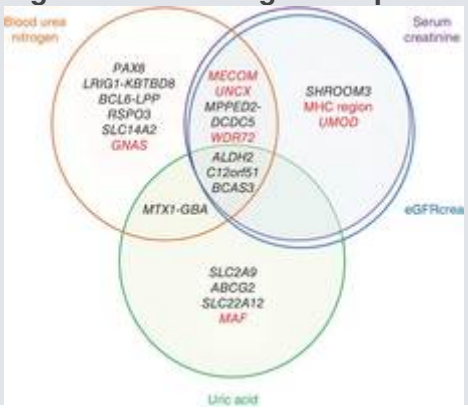
At a glance

1. **Figure 1: Manhattan plots of the GWAS meta-analysis for kidney function–related traits.**



Shown are the $-\log_{10}$ (P values) of the SNPs for the concentrations of blood urea nitrogen, serum creatinine and uric acid, and for eGFRcrea. The genetic loci that satisfied the genome-wide significance threshold of $P < 5.0 \times 10^{-8}$ (gray horizontal dotted line) in the combined study of the GWAS meta-analysis and replication are labeled for each of the traits. The newly identified loci are colored red, and the previously known loci are colored blue. The SNPs for which the P value was smaller than 1.0×10^{-20} are indicated at the upper limit of each plot.

2. **Figure 2: Venn diagram of pleiotropic associations of the identified loci.**



Genetic loci identified in the study are classified on the basis of the results of the pleiotropic association study of kidney function–related traits (Table 2 and Supplementary Table 8). Genes that showed significant associations with risk for stage 3+ CKD are colored red.

Main

Chronic kidney disease—the impairment of kidney function—constitutes a serious public health burden on society worldwide, with increased risks of mortality and morbidity[1, 2]. Biochemical measures of kidney function that are commonly used in clinical practice include the concentrations of blood urea nitrogen, serum creatinine and uric acid and glomerular

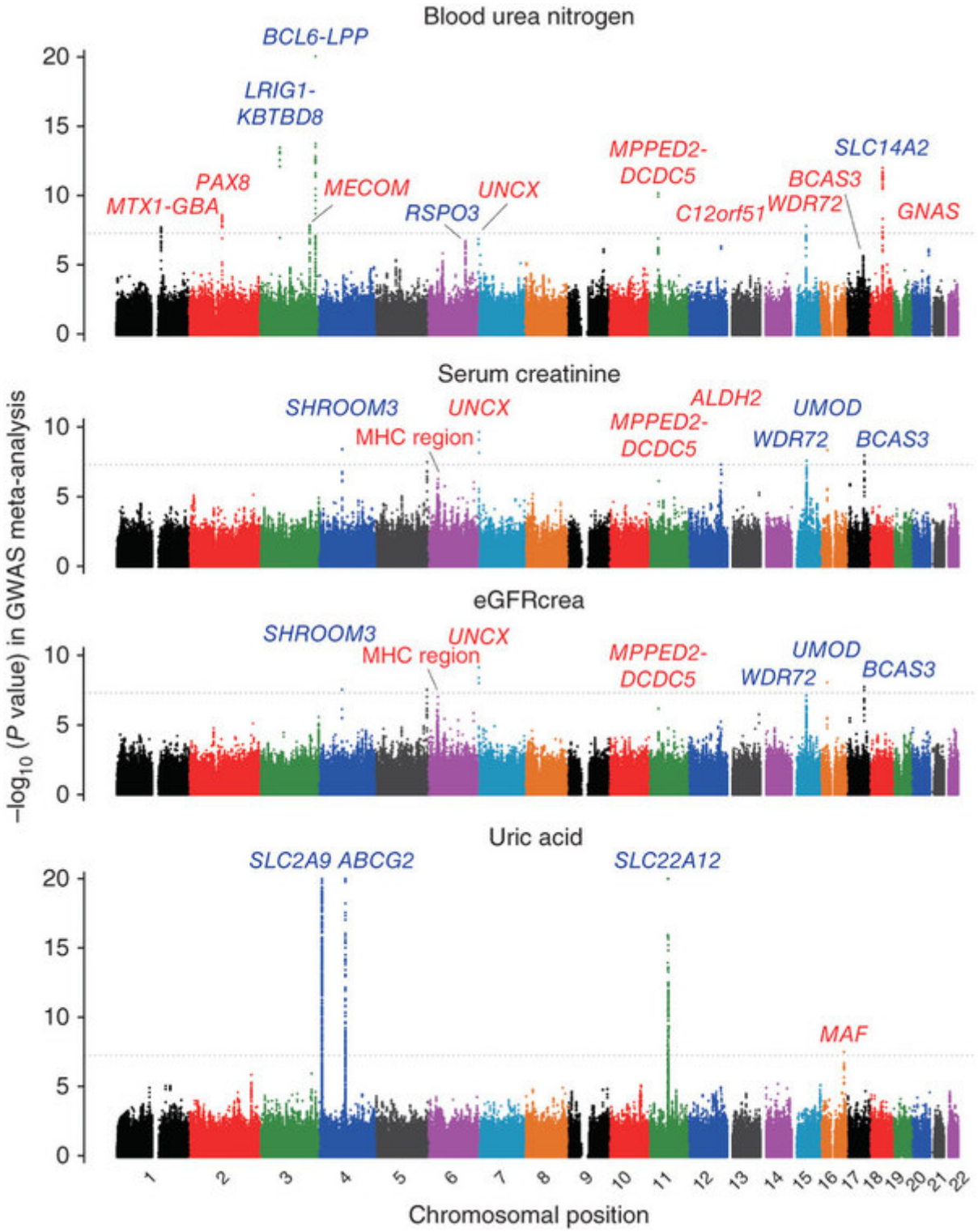
filtration rate (GFR). Heritability estimates have shown that genetic factors contribute significantly to interindividual variance in kidney function[3], and recent developments in genome-wide association studies (GWAS) have identified a number of genetic loci associated with measurements of kidney function[4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. However, most of these studies were conducted in populations of European ancestry[4, 6, 7, 8, 10, 11, 12, 13], and the extension of GWAS approaches to non-European populations would provide an opportunity to discover additional loci. We report a large-scale meta-analysis of GWAS and a replication study of kidney function–related traits involving 71,149 east Asian subjects performed by AGEN[9, 14, 15, 16] in which 11 cohorts participated (BBJ, SP2, SiMES, SINDI, SCES, KARE, HEXA, TWSC, TWT2D, GenSalt and CAGE; Online Methods and Supplementary Note).

In this study, we evaluated four kidney function–related traits (Supplementary Tables 1 and 2): the concentrations of blood urea nitrogen (n = 57,178), uric acid (n = 33,074) and serum creatinine (n = 61.919) and eGFRcrea (n = 62,087). Blood urea nitrogen concentration reflects the amount of nitrogen in the blood and is related to protein metabolism, including excretion by the kidneys[17]. Uric acid is the end product of purine metabolism, and impaired renal excretion of uric acid leads to hyperuricemia. Epidemiological studies suggest that uric acid is a risk factor for various diseases, including gout and myocardial infarction[13]. Serum creatinine levels and eGFRcrea are the most common kidney function measures used for the definition of CKD[1, 2], for which extensive genetic studies in European populations have been conducted[4, 6, 7, 8].

The GWAS meta-analysis included 51,327 east Asian individuals and evaluated approximately 2.4 million autosomal SNPs with a minor allele frequency (MAF) of ≥0.01. These SNPs were obtained by imputation of genotypes on the basis of HapMap Phase 2 panels (Supplementary Tables 3 and 4). The inflation factors of the test statistics were modest (λ_{GC} = 1.060, 1.072, 1.079 and 1.031 for blood urea nitrogen, serum creatinine, eGFRcrea and uric acid, respectively), which suggested that population structures did not have a substantial impact on the results of the meta-analysis. Quantile-quantile plots of the P values indicated notable discrepancies in their tails from those anticipated under the null hypothesis of no association, indicating the presence of significant associations in the meta-analysis (Supplementary Fig. 1). We identified 25 associations that satisfied the genome-wide significance threshold of $P < 5.0 \times 10^{-8}$. Of these, eight, seven, six and four genetic loci were found to associated with blood urea nitrogen, serum creatinine, eGFRcrea and uric acid, respectively (Supplementary Table 5).

We then performed an *in silico* replication study using data from an additional 19,822 east Asians for the loci that associated at $P < 5.0 \times 10^{-6}$ in the GWAS meta-analysis. Through the combined study of the GWAS meta-analysis and replication, we identified 32 significant associations at $P < 5.0 \times 10^{-8}$ (13, 8, 7 and 4 loci for blood urea nitrogen, serum creatinine, eGFRcrea and uric acid, respectively; Fig. 1 and Supplementary Table 5). We found that seven of these newly associated loci were associated with both serum creatinine concentration and eGFRcrea, with the same landmark SNPs involved at each locus, which reflects the close relationship between these two phenotypes ($R^2 = 0.76$ for common log-transformed values; Supplementary Table 6)[1, 2]. Among the loci identified in the combined analysis, associations at 15 loci were previously reported[4, 5, 6, 7, 8, 9, 10, 11, 12, 13]: *LRIG1-KBTBD8*, *BCL6-LPP*, *RSPO3* and *SLC14A2* for blood urea nitrogen concentration (smallest $P = 8.8 \times 10^{-30}$ at *BCL6-LPP*); *SHROOM3*, *WDR72*, *UMOD* and *BCAS3* for serum creatinine concentration (smallest $P = 1.2 \times 10^{-13}$ at *WDR72*); *SHROOM3*, *WDR72*, *UMOD* and *BCAS3* for eGFRcrea (smallest $P = 6.0 \times 10^{-13}$ at *WDR72*); and *SLC2A9*, *ABCG2* and *SLC22A12* for uric acid concentration (smallest $P = 1.6 \times 10^{-65}$ at *SLC2A9*). At the *UMOD* locus, the rs12917707 variant associated with eGFRcrea in Europeans[4, 7] had a low MAF (<0.01) and was not evaluated in our GWAS meta-analysis. However, we identified another variant that showed a significant association with eGFRcrea ($P = 3.6 \times 10^{-10}$ at rs11864909; MAF = 0.19; $r^2 = 0.02$ with rs12917707).

Figure 1: Manhattan plots of the GWAS meta-analysis for kidney function–related traits.



Shown are the $-\log_{10}$ (P values) of the SNPs for the concentrations of blood urea nitrogen, serum creatinine and uric acid, and for eGFRcrea. The genetic loci that satisfied the genome-wide significance threshold of $P < 5.0 \times 10^{-8}$ (gray horizontal dotted line) in the combined study of the GWAS meta-analysis and replication are labeled for each of the traits. The newly identified loci are colored red, and the previously known loci are colored blue. The

SNPs for which the *P* value was smaller than 1.0×10^{-20} are indicated at the upper limit of each plot.

In addition, we identified 17 loci newly associated with kidney function–related traits (Table 1 and Supplementary Fig. 2). Namely, we identified associations at nine loci for blood urea nitrogen concentration (*MTX1-GBA*, *PAX8*, *MECOM*, *UNCX*, *MPPED2-DCDC5*, *C12orf51*, *WDR72*, *BCAS3* and *GNAS* at 1q22, 2q13, 3q26, 7p22, 11p14, 12q24.13, 15q21, 17q23 and 20q13, respectively; smallest $P = 4.5 \times 10^{-16}$ at rs10767873 in *MPPED2-DCDC5*), four loci for serum creatinine concentration (the major histocompatibility (MHC) region, *UNCX*, *MPPED2-DCDC5* and *ALDH2* at 6p21, 7p22, 11p14 and 12q24.2, respectively; smallest $P = 4.6 \times 10^{-11}$ at rs10277115 in *UNCX*), three loci for eGFRcrea (the MHC region, *UNCX* and *MPPED2-DCDC5* at 6p21, 7p22 and 11p14, respectively; smallest $P = 1.0 \times 10^{-10}$ at rs10277115 in *UNCX*) and one locus for uric acid concentration (*MAF* at 16q23, $P = 1.1 \times 10^{-9}$ at rs889472). Combinations of these identified loci explained 1.3%, 0.54%, 0.55% and 2.3% of interindividual variance in blood urea nitrogen, serum creatinine, eGFRcrea and uric acid, respectively.

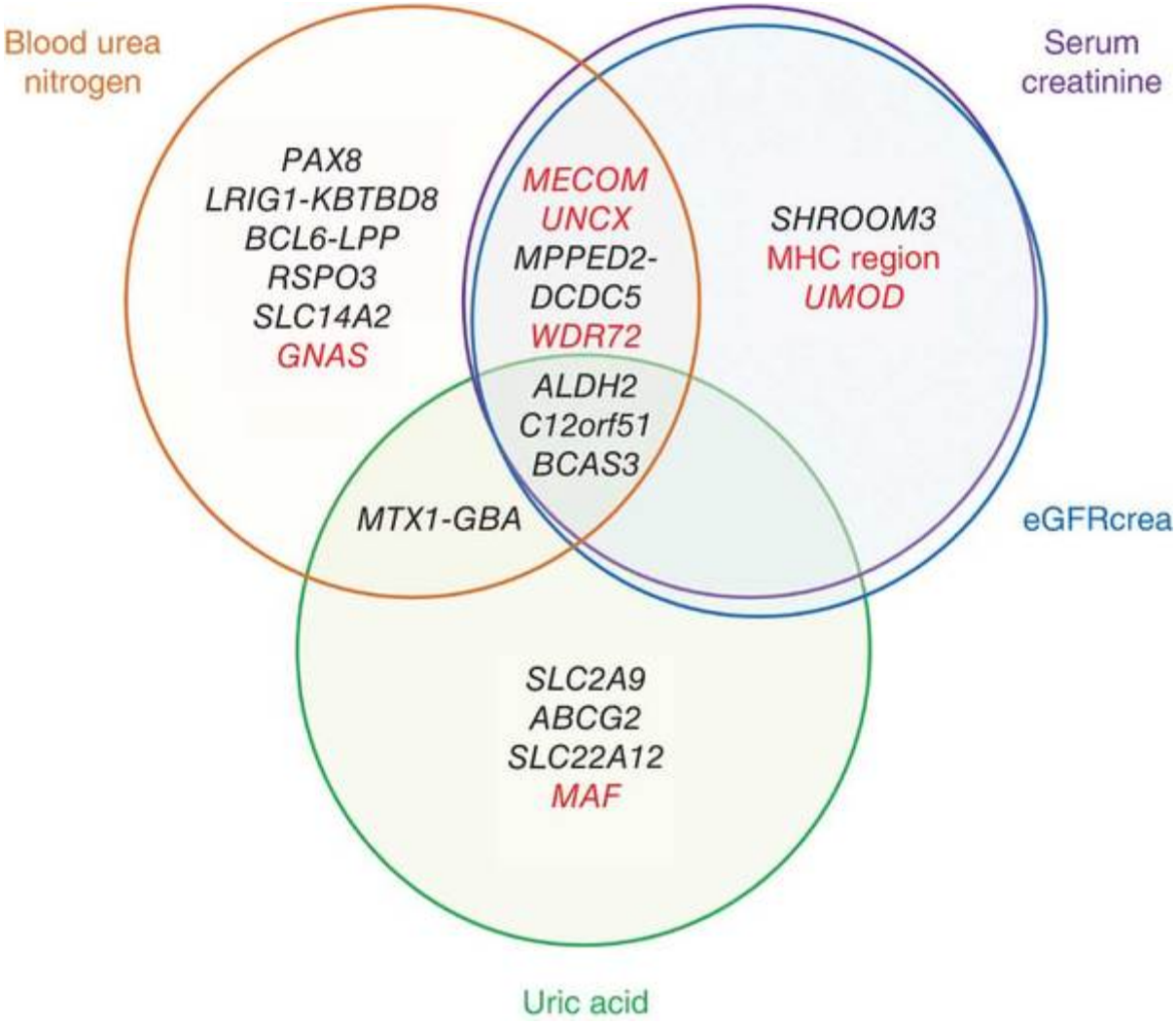
To determine whether the associations that we observed were relevant to populations of European ancestry, we evaluated the newly associated loci (at $P < 5.0 \times 10^{-8}$ in our combined meta-analysis) in Europeans by using the results of studies by the KidneyGen (*n* = 23,812 for serum creatinine concentration)[6], CKDGen (*n* = 67,093 for eGFRcrea)[7] and GUGC (*n* = 110,347 for uric acid concentration; A. Köttgen *et al.*, personal communication) consortia. Nine of the 15 loci that reached $P < 5.0 \times 10^{-8}$ for serum creatinine, eGFRcrea and uric acid measures in our study also showed significant associations in the European study ($P < 0.05/15 = 0.0033$, Bonferroni correction for the number of loci with available results), including the *MPPED2-DCDC5* locus for eGFRcrea ($P = 5.3 \times 10^{-8}$ at rs963837; Supplementary Table 5).

We also evaluated the loci previously reported to be associated with kidney function measures, after excluding the 14 loci that had already been identified in our study (Supplementary Table 7)[4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. Of the 31 loci evaluated, we replicated associations at 8 loci in our study ($P < 0.05/31 = 0.0016$, Bonferroni correction for the number of loci), including in *CPS1*, *RGS14*, *STC1*, *RNASEH2C-OVOL1* and *SLC6A13* for eGFRcrea and *GCKR*, *LRP2* and *LRRC16A-SLC17A1* for uric acid concentration.

As the evaluated phenotypes reflect both common and unique biological aspects of kidney status, it is of interest to understand whether the loci associated with kidney function traits

show pleiotropic patterns of associations[18]. We evaluated the associations of the identified loci within the evaluated kidney function–related traits and risk of stage 3+ CKD (defined as eGFRcrea of <60 ml/min/1.73 m^2; Fig. 2, Table 2 and Supplementary Table 8)[1,2]. Of 21 unique loci, 9 yielded significant associations with three or more phenotypes ($P < 0.05/21 = 0.0024$, Bonferroni correction for the number of loci). In particular, the *ALDH2*, *C12orf51* and *BCAS3* loci had significant associations with all of the evaluated kidney function–related traits. We also observed significant risk for CKD at several loci, including in *MECOM*, the MHC region, *UNCX*, *WDR72*, *UMOD*, *MAF* and *GNAS*. Because of the definition of CKD[1,2], previous studies assessed CKD risk primarily at the loci associated with serum creatinine concentration and eGFRcrea[4,6,7,8]. However, our results suggest that genetic risk for CKD would also be contributed to by other kidney function–related loci, such as *MAF* and *GNAS*. Recent studies suggested the superiority of eGFR based on serum cystatin C concentration (eGFRcys) relative to eGFRcrea, especially for predicting GFR in subjects with normal or mildly reduced GFR, and assessment of the genetic factors underlying eGFRcys in east Asians would thus be warranted.

Figure 2: Venn diagram of pleiotropic associations of the identified loci.



Genetic loci identified in the study are classified on the basis of the results of the pleiotropic association study of kidney function–related traits (Table 2 and Supplementary Table 8). Genes that showed significant associations with risk for stage 3+ CKD are colored red.

In this study, we identified new associations at *MTX1-GBA*, *PAX8*, *MECOM*, the MHC region, *UNCX*, *MPPED2-DCDC*5, *ALDH2*, *C12orf51*, *WDR72*, *MAF*, *BCAS3* and *GNAS* with kidney function–related traits. *MTX1* has an essential role in embryonic development, and *GBA* encodes glucocerebrosidase, an enzyme mediating glycolipid metabolism[19]. Both are known as causal genes in Gaucher disease[19], a lysosomal storage disease, although kidney function decline has not been implicated in pathogenesis. *PAX8* is a member of the PAX gene family and is widely expressed in renal tissues[20]. *MECOM* (also known as *EVI1*) encodes a transcriptional regulator involved in hematopoiesis[21]. The MHC region contains a large number of genes related to the immune system, including human leukocyte antigen (HLA) genes. The SNP that was found to be associated with serum creatinine concentration and eGFRcrea (rs3828890) was located in the MHC class I region[22] and was in moderate

linkage disequilibrium with the *HLA-DRB1*1302* and *HLA-DQB1*0604* alleles (D' > 0.65 and r^2 > 0.40 for both alleles)[23]. *UNCX* encodes a paired-type homeobox transcription factor that has essential roles in skeleton formation and kidney development[24]. The function of *MPPED2* is as yet unknown, and *DCDC5* encodes a protein with two doublecortin domains, which serve as protein-interaction platforms[25]. It is noteworthy that the *MTX1-GBA*, *MECOM* and *MPPED2-DCDC5* loci have been reported to influence serum magnesium levels[26], which are maintained by renal regulation of magnesium reabsorption. The loci in *ALDH2*, *WDR72* and *BCAS3* have been reported to be associated with some kidney function measures[5, 7], although the biological roles of these genes in renal homeostasis have not been substantially explored. Although the function of the protein encoded by *C12orf51* has not been examined, this locus was reported to be associated with serum lipid and liver enzyme concentrations in east Asians[9]. *MAF* encodes a leucine zipper transcription factor and has been implicated in the pathogenesis of minimal-change nephrotic syndrome (MCNS)[27]. Defects in *MAF* cause juvenile-onset pulverulent cataract as well as congenital cerulean cataract (CCA4)[28]. *GNAS* encodes the heterotrimeric G protein $G_s\alpha$, and the associated locus in this gene is also associated with multiple metabolic traits, including blood pressure, in Europeans[29]. Nevertheless, other genes near each of the loci could also be candidates, and further functional assessment is desirable.

In conclusion, in this large-scale meta-analysis in east Asian populations, we identified multiple loci newly associated with kidney function–related traits and pleiotropic associations. Our study should make an important contribution to the enhanced understanding of the genetic architecture of kidney function.

URLs.

International HapMap Project, http://hapmap.ncbi.nlm.nih.gov/; MACH and mach2qtl software, http://www.sph.umich.edu/csg/abecasis/MACH/index.html; IMPUTE and SNPTEST software, http://www.stats.ox.ac.uk/~marchini/software/gwas/gwas.html; BEAGLE software, http://faculty.washington.edu/browning/beagle/beagle.html; PLINK software, http://pngu.mgh.harvard.edu/~purcell/plink/; R statistical software, http://cran.r-project.org/; SNAP software, http://www.broadinstitute.org/mpg/snap/index.php.

Methods

Subjects.

The 71,149 subjects included in the GWAS meta-analysis for kidney function–related traits (n = 57,178, 61,919, 62,087 and 33,074 for blood urea nitrogen, eGFRcrea and uric acid,

respectively) were obtained from 18 studies conducted in the following 11 population-, hospital- or family-based cohorts of east Asian populations through the collaborations of AGEN[9, 14, 15, 16]: the BioBank Japan Project (BBJ), the Singapore Prospective Study Program (SP2), the Singapore Malay Eye Study (SiMES), the Singapore Indian Study (SINDI), the Singapore Chinese Eye Study (SCES), the Korea Association Resource project (KARE), the Health Examinee shared control study (HEXA), the Taiwan Super Control Study (TWSC), the Taiwan Type 2 Diabetes Consortium (TWT2D), the Genetic Epidemiology Network of Salt Sensitivity (GenSalt) and Cardio-metabolic Genome Epidemiology (CAGE). Of these, 51,327 subjects were enrolled in the GWAS meta-analysis, and 19,822 subjects were enrolled in the *in silico* replication study. Some of the subjects were included in previous studies of east Asian populations[9, 14, 15, 16]. All participants in each cohort provided written informed consent for participation in the study, as approved by the ethical committees of each of the institutional review boards. Each study established a consensus on subject participation and phenotype definition and analytical protocol for the project. Detailed descriptions of the participating cohorts and the characteristics of the subjects are provided in Supplementary Tables 1 and 2 and in the Supplementary Note. Details of the European studies enrolled by the KidneyGen (n = 23,812 for serum creatinine concentration), CKDGen (n = 67,093 for eGFRcrea) and GUGC (n = 110,347 for uric acid concentration) consortia, including subject details and the study designs, have been described at length elsewhere (refs 6,7 and A. Köttgen *et al.*, personal communication).

Genotyping and quality control.

Genotyping platforms and quality control criteria, including exclusion of closely related subjects and outliers in terms of ancestry and cutoff values for sample call rate, SNP call rate, MAF and Hardy-Weinberg equilibrium P value are provided for each study (Supplementary Table 3 and Supplementary Note). Genotype imputation was performed on the basis of the HapMap Phase 2 panels (Japanese in Tokyo, Japan (JPT) and Han Chinese in Beijing, China (CHB) populations, except for SiMES and SINDI, for which JPT, CHB, Yoruba in Ibadan, Nigeria (YRI) and Utah residents of Northern and Western European ancestry (CEU) populations were adopted) by using MACH, IMPUTE or BEAGLE software (see URLs). After imputation, we excluded SNPs with MAF of <0.01 or imputation quality score of R^2 of <0.5 from each study.

Phenotype modeling.

Clinical information on the subjects, including age, gender and mean ± s.d. values for the kidney function–related traits, are provided (Supplementary Table 2). Collection methods for the clinical information in each of the cohorts are described (Supplementary Note). In this study, eGFRcrea was estimated on the basis of serum creatinine levels, using the Japanese coefficient-modified CKD Epidemiology Collaboration (CKD-EPI) equation[2]. We excluded subjects who were <18 or >85 years old, those who had eGFRcrea of <15 ml/min/1.73 m^2 and those who had undergone renal replacement therapy. Subjects with gastrointestinal bleeding, systemic infection or hepatic failure and subjects who had undergone uric acid–lowering therapy (alloprinol, benzbromarone or probenecid) were also excluded from the analyses for blood urea nitrogen and uric acid concentration, respectively.

Genome-wide association study.

Associations of SNPs with common log-transformed values of blood urea nitrogen (mg/dl), serum creatinine (mg/dl), eGFRcrea (ml/min/1.73 m^2) or non-transformed values of uric acid concentration (mg/dl) were assessed by linear regression models assuming additive effects of the allele dosages of the SNPs using mach2qtl, SNPTEST, PLINK or R statistical software (see URLs). For the subjects in the family-based cohort, generalized linear mixed models accounting for the family structure were applied. In the regression model, gender, age, drinking status (current drinker or not), smoking status (previous or current smoker or not), body mass index and other cohort-specific variables were incorporated as covariates (Supplementary Note).

GWAS meta-analysis.

In the GWAS meta-analysis, we included autosomal SNPs that satisfied quality control criteria in three or more GWAS for each of the traits, which yielded between 2.2 and 2.4 million SNPs (Supplementary Table 4). Information about the SNPs, including the coded alleles, was oriented to the forward strand of the NCBI Build 36 reference sequence. GWAS meta-analysis was performed using an inverse variance–weighted method, assuming a fixed-effects model for study-specific effect estimates (β) and standard errors (SE) of the coded alleles of the SNPs, using a Java source code implemented by the authors[30, 31]. Genomic control corrections were carried out on test statistics from each of the GWAS using study-specific inflation factors (λ_{GC}) and were applied again to the results of the GWAS meta-analysis (Supplementary Fig. 1)[32].

***In silico* replication study.**

The *in silico* replication study was conducted using additional independent east Asian subjects (Supplementary Tables 1 and 2) for the loci that satisfied $P < 5.0 \times 10^{-6}$ in the GWAS meta-analysis for each of the traits (17, 14, 14 and 6 loci for blood urea nitrogen, serum creatinine, eGFRcrea and uric acid, respectively; Supplementary Table 5). For each of the loci, the SNP that showed the most significant association was selected. The associations of the SNPs were assessed in the same manner as in the GWAS. The combined study of the GWAS meta-analysis and replication was conducted using an inverse variance method, assuming a fixed-effects model[30, 31]. The SNPs that satisfied $P < 5.0 \times 10^{-8}$ in the combined study were considered to be significantly associated with the relevant kidney function–related trait, and the associations of these SNPs were further evaluated using data in European populations from the KidneyGen, CKDGen and GUGC consortia (refs 6,7 and A. Köttgen *et al.*, personal communication).

Estimation of explained variance.

The interindividual variance in kidney function–related traits explained by the combination of the identified loci ($P < 5.0 \times 10^{-8}$ for each phenotype) was estimated using a genetic risk score model. We calculated the scores of the subjects enrolled in the *in silico* replication study by the BioBank Japan Project[33] (BBJ_5 and BBJ_6; Supplementary Table 2) by summing the dosages of the effect alleles carried by the subjects, which were weighted by the effect sizes of the SNPs obtained from the GWAS meta-analysis. The explained variance was estimated from linear regression models on the covariate-adjusted phenotypes by the scores.

Pleiotropic association analysis for kidney function–related phenotypes.

For the genetic loci that showed associations at $P < 5.0 \times 10^{-8}$ in the combined study, pleiotropic associations with the kidney function–related traits and with risk for stage 3+ CKD (defined as eGFRcrea of <60 ml/min/1.73m^2)[1, 2] were assessed. Associations with CKD risk were assessed using logistic regression models, incorporating the covariates using the subjects obtained from the BioBank Japan Project[33] (BBJ_1– BBJ_6; Supplementary Table 2).

References

1. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am. J. Kidney Dis.* **39**, S1–S266 (2002).

2. Horio, M., Imai, E., Yasuda, Y., Watanabe, T. & Matsuo, S. Modification of the CKD epidemiology collaboration (CKD-EPI) equation for Japanese: accuracy and use for population estimates. *Am. J. Kidney Dis.* **56**, 32–38 (2010).

3. Fox, C.S. *et al*. Genome-wide linkage analysis to serum creatinine, GFR, and creatinine clearance in a community-based population: the Framingham Heart Study. *J. Am. Soc. Nephrol.* **15**, 2457–2461 (2004).

4. Köttgen, A. *et al*. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat. Genet.* **41**, 712–717 (2009).

5. Kamatani, Y. *et al*. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat. Genet.* **42**, 210–215 (2010).

6. Chambers, J.C. *et al*. Genetic loci influencing kidney function and chronic kidney disease. *Nat. Genet.* **42**, 373–375 (2010).

7. Köttgen, A. *et al*. New loci associated with kidney function and chronic kidney disease. *Nat. Genet.* **42**, 376–384 (2010).

8. Gudbjartsson, D.F. *et al*. Association of variants at *UMOD* with chronic kidney disease and kidney stones—role of age and comorbid diseases. *PLoS Genet.* **6**, e1001039 (2010).

9. Kim, Y.J. *et al*. Large-scale genome-wide association studies in east Asians identify new genetic loci influencing metabolic traits. *Nat. Genet.* **43**, 990–995 (2011).

10. Döring, A. *et al*. SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. *Nat. Genet.* **40**, 430–436 (2008).

11. Vitart, V. *et al*. SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nat. Genet.* **40**, 437–442 (2008).

12. Kolz, M. *et al*. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet.* **5**, e1000504 (2009).

13. Yang, Q. *et al*. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ. Cardiovasc. Genet.* **3**, 523–530 (2010).

14. Kato, N. *et al*. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat. Genet.* **43**, 531–538 (2011).

0.

15. Okada, Y. *et al*. Common variants at *CDKAL1* and *KLF9* are associated with body mass index in east Asian populations. *Nat. Genet.* **44**, 302–306 (2012).

16. Wen, W. *et al*. Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.* **44**, 307–311 (2012).

17. Lyman, J.L. Blood urea nitrogen and creatinine. *Emerg. Med. Clin. North Am.* **4**, 223–233 (1986).

18. Okada, Y. *et al*. Genome-wide association study for C-reactive protein levels identified pleiotropic associations in the *IL6* locus. *Hum. Mol. Genet.* **20**, 1224–1231 (2011).

19. LaMarca, M.E. *et al*. A novel alteration in metaxin 1, F202L, is associated with N370S in Gaucher disease. *J. Hum. Genet.* **49**, 220–222 (2004).

20. Tong, G.X. *et al*. Expression of *PAX8* in normal and neoplastic renal tissues: an immunohistochemical study. *Mod. Pathol.* **22**, 1218–1227 (2009).

21. Kataoka, K. *et al*. Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. *J. Exp. Med.* **208**, 2403–2416 (2011).

22. MHC Sequencing Consortium. Complete sequence and gene map of a human major histocompatibility complex. *Nature* **401**, 921–923 (1999).

23. de Bakker, P.I. *et al*. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* **38**, 1166–1172 (2006).

24. Mansouri, A. *et al*. Paired-related murine homeobox gene expressed in the developing sclerotome, kidney, and nervous system. *Dev. Dyn.* **210**, 53–65 (1997).

25. Reiner, O. *et al*. The evolving doublecortin (DCX) superfamily. *BMC Genomics* **7**, 188 (2006).

26. Meyer, T.E. *et al*. Genome-wide association studies of serum magnesium, potassium, and sodium concentrations identify six loci influencing serum magnesium levels. *PLoS Genet.* **6**, e1001045 (2010).

27. Sahali, D. *et al*. A novel approach to investigation of the pathogenesis of active minimal-change nephrotic syndrome using subtracted cDNA library screening. *J. Am. Soc. Nephrol.* **13**, 1238–1247 (2002).

28. Vanita, V., Singh, D., Robinson, P.N., Sperling, K. & Singh, J.R. A novel mutation in the DNA-binding domain of *MAF* at 16q23.1 associated with autosomal dominant “cerulean cataract” in an Indian family. *Am. J. Med. Genet. A* **140**, 558–566 (2006).

29. Ehret, G.B. *et al*. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).

30. Okada, Y. *et al*. Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. *PLoS Genet.* **7**, e1002067 (2011).

31. Okada, Y. *et al*. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat. Genet.* **44**, 511–516 (2012).

32. de Bakker, P.I. *et al*. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122–R128 (2008).

33. Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).

Acknowledgments

The authors acknowledge the essential roles of AGEN in developing the study. BBJ was supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT). SP2 was funded by grants from the Biomedical Research Council of Singapore (BMRC 05/1/36/19/413 and 03/1/27/18/216) and the National Medical Research Council of Singapore (NMRC/1174/2008). SiMES was funded by the National Medical Research Council of Singapore (NMRC 0796/2003, IRG07nov013 and NMRC/STaR/0003/2008) and the Biomedical Research Council of Singapore (BMRC 09/1/35/19/616). SINDI and SCES were funded by grants from the Biomedical Research Council of Singapore (BMRC 09/1/35/19/616 and BMRC 08/1/35/19/550) and the National Medical Research Council of Singapore (NMRC/STaR/0003/2008). Y.-Y.T. acknowledges support from the Singapore National Research Foundation (NRF-RF-2010-05). E.-S.T. receives support from the National Medical Research Council of Singapore through a Clinician Scientist Award. We thank the Singapore BioBank and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, for providing services for tissue archiving and genotyping, respectively. KARE was supported by grants from the Korea Centers for Disease Control and Prevention (4845-301, 4851-302 and 4851-307) and an intramural grant from the Korea National Institute of Health (2010-N73002-00). Y.S.C. acknowledges support from a National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (2012R1A2A1A03006155). TWSC and TWT2D were supported by the Academia Sinica Genomic Medicine Multicenter Study (40-05-GMM). We acknowledge the National Center for Genome Medicine (NSC100-2319-B-001-001), the National Core Facility Program for Biotechnology of the National Science Council, Taiwan, for technical help in sample

management and genotyping. GenSalt was supported by grants (U01HL072507, R01HL087263 and R01HL090682) from the National Heart, Lung, and Blood Institute, the US National Institutes of Health. CAGE was supported by grants for Core Research for Evolutional Science and Technology (CREST) from the Japan Science Technology Agency; the Program for Promotion of Fundamental Studies in Health Sciences, the National Institute of Biomedical Innovation Organization (NIBIO) and the grant of National Center for Global Health and Medicine (NCGM). We thank all the people who supported the Hospital-based Cohort Study at NCGM and the Amagasaki Study. We thank A. Taniguchi, H. Rakugi, K. Sugimoto, K. Kamide and C. Makibayashi for supporting the study.

Author information

Primary authors

1. **These authors contributed equally to this work.**

- Yukinori Okada,
- Xueling Sim,
- Min Jin Go,
- Jer-Yuarn Wu,
- Dongfeng Gu &
- Fumihiko Takeuchi

Affiliations

1. **Laboratory for Statistical Analysis, Center for Genomic Medicine (CGM), RIKEN, Yokohama, Japan.**

- Yukinori Okada &
- Atsushi Takahashi

2. **Department of Allergy and Rheumatology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan.**

- Yukinori Okada &
- Kazuhiko Yamamoto

3. **Centre for Molecular Epidemiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.**

- Xueling Sim &
- Yik-Ying Teo

4. **Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA.**

- Xueling Sim

5. **Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Republic of Korea.**

- Min Jin Go,
- Young Jin Kim,
- Jong-Young Lee,
- Bok-Ghee Han &
- Yoon Shin Cho

6. **Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.**

- Jer-Yuarn Wu,
- Chien-Hsiun Chen,
- Li-Ching Chang,
- S-J Cathy Fann &
- Yuan-Tsong Chen

7. **School of Chinese Medicine, China Medical University, Taichung, Taiwan.**

- Jer-Yuarn Wu,
- Chien-Hsiun Chen &
- Fuu-Jen Tsai

8. **State Key Laboratory of Cardiovascular Diseases, Fu Wai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.**

- Dongfeng Gu &

- Shufeng Chen

9. **Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan.**

- Fumihiko Takeuchi,
- Masato Isono &
- Norihiro Kato

10. **Laboratory for Endocrinology and Metabolism, CGM, RIKEN, Yokohama, Japan.**

- Shiro Maeda

11. **Laboratory for Medical Informatics, CGM, RIKEN, Yokohama, Japan.**

- Tatsuhiko Tsunoda

12. **Saw Swee Hock School of Public Health, National University of Singapore, Singapore.**

- Peng Chen,
- Su-Chi Lim,
- Yik-Ying Teo &
- E-Shyong Tai

13. **Diabetes Centre, Khoo Teck Puat Hospital, Singapore.**

- Su-Chi Lim

14. **Department of Medicine, Khoo Teck Puat Hospital, Singapore.**

- Su-Chi Lim

15. **Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.**

- Tien-Yin Wong &
- Tin Aung

16. **Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.**

- Tien-Yin Wong &
- Tin Aung

17. **Centre for Eye Research Australia, University of Melbourne, East Melbourne, Victoria, Australia.**

- Tien-Yin Wong

18. **Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore.**

- Jianjun Liu &
- Yik-Ying Teo

19. **Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, USA.**

- Terri L Young

20. **Institute of Human Genetics, University of California, San Francisco, California, USA.**

- Mark Seielstad

21. **Department of Statistics and Applied Probability, National University of Singapore, Singapore.**

- Yik-Ying Teo

22. **National University of Singapore Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore.**

- Yik-Ying Teo

23. **Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea.**

- Daehee Kang

24. **Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana, USA.**

- Hao Mei &

- Jiang He

25. **Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri, USA.**

- Dabeeru C Rao

26. **Human Genetics Center, University of Texas School of Public Health, Houston, Texas, USA.**

- James E Hixson

27. **Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, Japan.**

- Tomohiro Katsuya

28. **Department of Geriatric Medicine and Nephrology, Osaka University Graduate School of Medicine, Suita, Japan.**

- Tomohiro Katsuya &
- Toshio Ogihara

29. **Morinomiya University of Medical Sciences, Osaka, Japan.**

- Toshio Ogihara

30. **Department of Epidemiology and Biostatistics, Imperial College London, London, UK.**

- John C Chambers &
- Weihua Zhang

31. **National Heart and Lung Institute, Imperial College London, London, UK.**

- Jaspal S Kooner

32. **A full list of contributing members and affiliations is provided in the** Supplementary Note**.**

- $affiliationAuthor

33. **Institute of Genetic Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany.**

- Eva Albrecht

34. **Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan.**

- Michiaki Kubo

35. **Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan.**

- Yusuke Nakamura

36. **Laboratory for International Alliance, CGM, RIKEN, Yokohama, Japan.**

- Naoyuki Kamatani

37. **Department of Biomedical Science, Hallym University, Gangwon-do, Republic of Korea.**

- Yoon Shin Cho

38. **Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.**

- E-Shyong Tai

39. **Duke–National University of Singapore Graduate Medical School, Singapore.**

- E-Shyong Tai

40. **Laboratory for Cardiovascular Diseases, CGM, RIKEN, Yokohama, Japan.**

- Toshihiro Tanaka

41. **These authors jointly directed this work.**

- Norihiro Kato,
- Jiang He,
- Yuan-Tsong Chen,
- Yoon Shin Cho,

- E-Shyong Tai &
- Toshihiro Tanaka

Consortia

1. **The KidneyGen Consortium**

2. A full list of contributing members and affiliations is provided in the Supplementary Note.

3. **The CKDGen Consortium**

4. A full list of contributing members and affiliations is provided in the Supplementary Note.

5. **The GUGC consortium**

6. A full list of contributing members and affiliations is provided in the Supplementary Note.

Contributions

Y.O. and T. Tanaka designed the overall study. Y.O., X.S., M.J.G., C.-H.C., D.G., F.T. and P.C. analyzed GWAS data. Y.O. performed meta-analysis and other statistical analysis. Y.O., A.T., S.M., T. Tsunoda, K.Y., M.K., Y.N., N. Kamatani and T. Tanaka managed GWAS data of BBJ. X.S., P.C., S.-C.L., T.-Y.W., J.L., T.L.Y., T.A., M.S., Y.-Y.T. and E.-S.T. managed the GWAS data from SP2, SiMES, SINDI and SCES. M.J.G., Y.J.K., J.-Y.L., B.-G.H., D.K. and Y.S.C. managed the GWAS data from KARE and HEXA. C.-H.C., F.-J.T., L.-C.C., S.-J.C.F., Y.-T.C. and J.-Y.W. managed the GWAS data from TWSC and TWT2D. D.G., H.M., D.C.R., J.E.H., S.C. and J.H. managed the GWAS data from GenSalt. F.T., T.K., M.I., T.O. and N. Kato managed the GWAS data from CAGE. J.C.C., W.Z. and J.S.K. managed the data from the KidneyGen Consortium. E.A. managed the data from the GUGC consortium. Y.O., T. Tanaka, E.-S.T., Y.S.C., J.-Y.W., J.H. and N. Kato directed the study and wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.